The multifactorial etiology of celiac disease explored by combining several national registers

Fredinah Namatovu

Department of Public Health and Clinical Medicine
Epidemiology and Global Health
Umeå University 2015
Dedicated

To my Parents
# Table of Contents

**Abstract** iii  
**Original papers** v  
**List of Abbreviations** vi  
**Background** 1  
- Celiac Disease 1  
- Celiac Disease in history 1  
- Global overview 2  
- Geographical variation in Celiac Disease incidence 3  
- Celiac Disease in Sweden 3  
- Celiac Disease diagnosis 4  
  - The multifactorial etiology of Celiac Disease 4  
**Aims of the thesis** 9  
**Conceptual frameworks** 11  
**Methods** 13  
- Overview of the study designs 13  
- Study setting 13  
- Data source 14  
- Study variables and statistical analysis 16  
**Results** 21  
- Incidence rate 21  
- Geographical variation in Celiac Disease 23  
- Environmental factors associated with Celiac Disease 27  
**Discussion** 29  
- Main findings 29  
- Strengths and limitations 29  
- Celiac Disease incidence is increasing 31  
- Geographical variation in Celiac Disease risk 31  
- The life course perspective and Celiac Disease risk factors 32  
- The sufficient-component cause model and Celiac Disease risk factors 39  
**Concluding remarks and future research** 41  
**The researcher** 43  
**Acknowledgements** 45  
**References** 47  
**Appendix** 61
Abstract

Background
Celiac Disease (CD) is a systemic disease with chronic small intestinal immune-mediated enteropathy occurring in genetically predisposed individuals. CD is triggered by dietary gluten found in wheat, rye, and barley and is considered a public health problem.

Objective
The aim of this thesis was to estimate CD incidence in Sweden and to investigate environmental and lifestyle factors that might influence the risk of developing CD during childhood, which might guide future approaches to CD prevention.

Methods
A quantitative approach was used to analyze data obtained from Swedish national registers accessed through the Umeå SIMSAM Lab. CD cases aged 0–14.9 years at diagnosis were obtained through the National Swedish Childhood CD Register. Data on total child population with demographic and socioeconomic conditions were provided by Statistics Sweden, and health data were from the National Board of Health and Welfare. Data linkage was performed by Statistics Sweden using the personal identity number (PIN). During the study duration of 1973 to 2009, we identified 9 107 children with CD, out of these data linkage was possible for 6 569 cases reported with a PIN. The total number of live births were 1 578 094, and there were 28 039 741 person-years of follow-up.

Results
During the follow-up period from 1973 to 2009, the CD incidence among all ages (0–14.9 years) gradually increased, aside from a temporary decline in 1996–1997. At the end of follow-up, a significant decline in the incidence of CD among those aged <2 years occurred, while children aged 2–14.9 years experienced a continuing gradual increase. The median age at diagnosis increased from 1.0 years of age in the 1970s to 6.8 years of age in 2009. CD risk was much higher in children born in southern Sweden (adjusted hazard ratio [HR] 1.59; 95% confidence interval [CI] 1.48–1.72) compared to those in the north. Variation in CD incidence was more pronounced on the municipality level compared to the regional level. On the neighborhood level, we found a positive association between work income and CD risk (odds ratio [OR] 2.06; 95% CI 1.61–2.64), but the risk was significantly reduced by increased average age and in areas with high level of industrial and
commercial activity. On the individual level, a reduction in risk was observed in children whose mothers were aged ≥35 years (OR 0.8; 95% CI 0.7–0.9) and with high maternal income (OR 0.9; 95% CI 0.8–0.9). Being a second-born child was positively associated with CD. Among boys, elective caesarean delivery increased CD risk (OR 1.2; 95% CI 1.0–1.4), while maternal overweight, premature rupture of the membrane, and low birth weight all showed negative associations. In girls, the risk of developing CD was associated with repeated maternal urinary tract infections during pregnancy. With regard to season of birth, we found an increased CD risk in children aged <2 years if they were born during the spring, while those aged 2–14.9 years had a more prominent risk if they were born during summer or autumn.

Conclusion
Our findings reveal an increased CD incidence over time and an increase in age at diagnosis. CD risk varied between birth cohorts and with season of birth, suggesting cyclic environmental and/or lifestyle risk factors in CD etiology. Neighborhood composition influenced CD incidence, and this is one of the first attempts at identifying factors associated with the geographical variation of CD. The observed effect of elective caesarean delivery, repeated maternal urinary tract infections, and season of birth might be due to infections and/or unfavorable colonization of the microbiota during early life. Reduced CD risk associated with high maternal age and income might be due to other lifestyle conditions. More research on underlying risk factors for CD is required in order to develop more effective preventive strategies.
Original papers

The thesis is based on the papers listed below and referred to as Papers I–IV.

I. Fredinah Namatovu, Olof Sandström, Cecilia Olsson, Marie Lindkvist, Anneli Ivarsson. Celiac disease risk varies between birth cohorts, generating hypotheses about causality: evidence from 36 years of population-based follow-up. BMC Gastroenterol. 2014;2;14:59.


III. Fredinah Namatovu, Marie Lindkvist, Cecilia Olsson, Anneli Ivarsson, Olof Sandström. Maternal and perinatal conditions and the risk of developing celiac disease during childhood. (Submitted).

IV. Fredinah Namatovu, Marie Lindkvist, Cecilia Olsson, Anneli Ivarsson, Olof Sandström. Season and region as risk factors for celiac disease: a key to the etiology? (Submitted).

Paper I and II are reprinted with permission from the BioMedCentral, Copyright © 2015 BioMed Central Ltd and SAGE publications, respectively.
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD</td>
<td>Celiac disease</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>ESPGHAN</td>
<td>European Society for Pediatric Gastroenterology Hepatology and Nutrition</td>
</tr>
<tr>
<td>HR</td>
<td>Hazard ratio</td>
</tr>
<tr>
<td>NUTS</td>
<td>Nomenclature of territorial units for statistics</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>PIN</td>
<td>Personal identification number</td>
</tr>
<tr>
<td>PROM</td>
<td>Premature rupture of the membrane</td>
</tr>
<tr>
<td>SAMS</td>
<td>Small area market statistics</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Background

Celiac Disease

Celiac Disease (CD) is a chronic autoimmune small intestinal immune-mediated enteropathy triggered by ingestion of wheat gluten and related proteins from barley and rye among genetically predisposed individuals [1]. The clinical presentation of CD varies between individuals and even within an individual over time and can be diagnosed at any age [2-4].

Classical CD clinical manifestations during childhood are related to gastrointestinal and nutrient malabsorption, and these are exemplified by growth failure, diarrhea, abdominal distention, and lethargy [3-5]. However, the clinical picture of CD has changed over time, and these classical symptoms have become less dominant [3]. Currently, it is becoming more common for CD patients to present a non-classical form of CD with extra-intestinal features like osteoporosis, short stature, infertility, anemia, and neurological problems. It is important to note that in some cases signs and symptoms are subtle or even absent, thus CD often goes unrecognized causing many CD cases to be missed [2-7].

Adhering to a strict life-long gluten-free diet is the only effective treatment for CD because this helps restore the mucosa of the small intestine and relieves most of the signs and symptoms [7-15]. It should be noted, however, that gluten is a very common ingredient in the human diet, thus its total elimination presents a significant challenge for CD patients and their caregivers. Gluten-free products are not often available, and when they are they tend to be more expensive than ordinary foods thus making compliance to a gluten-free diet cumbersome [16].

Celiac Disease in history

A condition that resembles CD first became evident soon after the establishment of civilization and domestication that came hand in hand with the agricultural revolution [17, 18]. Agriculture led to the introduction of wheat and barley, which were the first cereals to be domesticated and thus the first to be introduced into the human diet. The majority of mankind adapted to these dietary changes, but unfortunately some did not and CD came into existence [17, 18]. The earliest description of a condition with CD-like symptoms is traced as far back as the first and second century AD in the works
of a Greek physician named Aretaeus of Cappadocia. He provided a
description of malabsorptive diarrhea, a symptom considered to be one of the
common manifestations of CD [17, 19].

More knowledge of CD was provided in the 19th century by Dr. Mathew Baillie
who described a CD-like condition characterized by chronic diarrhea in adults
causing malnutrition and a gas-distended abdomen [18]. Full credit for the
first modern description of CD is attributed to the English pediatrician
Samuel Jones Gee. In 1888 he described “The coeliac affection” as a condition
common in people of all ages. Patients often presented with chronic
indigestion, wasting, weakness, and diarrhea. He also hypothesized that this
condition could be cured through diet [19]. The progress in understanding CD
pathogenesis was further strengthened by major evidence provided in the
1940s by the Dutch pediatrician William Dicke when he linked CD etiology to
gluten. Dicke came to this conclusion during World War II when he observed
children with CD improve when they lacked access to wheat, rye and barley
thus identifying gluten as the major culprit [20]. Despite this early knowledge,
it was not until the 1950s that the effectiveness of a gluten-free diet in CD
treatment was recognized [18].

Another major milestone in our modern day understanding of CD was
reached in the 1950s by Margot Shiner who described a new jejunal biopsy
apparatus that she successfully used to biopsy the distal duodenum [18, 19,
21]. This advancement was followed by the discovery of a less cumbersome
and user-friendly capsule developed by Lieutenant Colonel William Holmes
Crosby that allowed doctors to link the disease to vivid patterns of damage to
the proximal small intestinal mucosa, and this led to the unraveling of the
mystery that surrounded CD pathogenesis [18, 19].

Evidence continued to build up, and by the 1980s the role of antibodies in CD
etiology was suggested and proven [17, 19, 22]. It was revealed that antibodies
like anti-gliadin, anti-endomysium, and the autoantibody anti-
transglutaminase can be found in the blood of children as a result of the
ingestion of gluten [17, 19]. This evidence led to the acceptance of CD as an
autoimmune disease triggered by the ingestion of gluten. Additionally, CD
was also linked specifically to the expression of the HLA DQ2 and HLA DQ8
genes [19, 22].

Global overview

Previously, CD was considered rare and was believed to only affect people of
European origin [20]. In recent years, however, evidence from Asia, Africa,
and Latin America has confirmed that CD is equally prevalent in these areas and as such CD has emerged as a global health condition [23-34]. Globally, the prevalence of CD is estimated to be between 0.5% and 1% [20, 23-34]. In the Saharawian children in an Algerian refugee camp, CD prevalence was estimated at 5.6% (although the uniqueness of this population might be attributed to their extreme living conditions) [25]. The diversity of CD has been further supported by screening studies and by the emergence of highly sensitive and specific CD serological markers that have revealed the often-missed subclinical forms of CD [35]. More and more evidence is pointing to CD as a global phenomenon, and thus one of the most common lifelong food-related diseases affecting mankind [23-34].

**Geographical variation in Celiac Disease incidence**

Geographical variation in CD has been reported in several populations [36]. CD incidence was found to be higher in Finland compared to the adjacent territory of Russian Karelia (1.4% versus 0.6%), even though the two populations have a similar genetic background [37]. Large regional variations in CD prevalence are also present in the United Kingdom [38] and in Sweden [39-41]. The reasons for these geographical variations in CD incidence are not yet clearly understood.

**Celiac Disease in Sweden**

Swedish studies on the child population in the mid-1960s and early 1970s reported CD incidences ranging between 0.16 and 1.43 per 1,000 live births [42-44]. By building up the National Swedish Childhood CD Register, our research group has followed CD incidence in the childhood population aged 0–15 years since 1973, with national coverage from 1998 [45]. The incidence rate was fairly stable between 1973 and 1984 at 10 cases per 100,000 person-years. The incidence rate started to increase rapidly from 1985 to 1994. This period was marked by a sharp increase that reached its peak in 1994 with 45 cases per 100,000 person-years. The incidence rate began to decrease in 1995, and by 1997 the epidemic had completely abated with an incidence rate of 16 cases per 100,000 person-years [45]. The follow-up period from 1998 to 2003 revealed that the annual incidence rate for all children younger than 15 years of age increased two-fold from 19 cases per 100,000 to 44 cases per 100,000 person-years [46].

The sudden rise in the incidence among children younger than 2 years during the period of 1985 to 1996 is an interesting phenomenon because such rates
Celiac Disease diagnosis

The 2012 European Society of Pediatrics Gastroenterology, Hepatology and Nutrition Association (ESPGHAN) suggested that CD diagnosis should be based on analysis of CD serological markers, HLA-DQ2 and HLA-DQ8 genotyping, an initial biopsy from the small intestine, and clinical and serological follow-up. As of September 2015, the National Institute for Health and Care Excellence (NICE) has published new recommendations for serological testing for CD that include total immunoglobulin A (IgA) and IgA tissue transglutaminase as a first choice for serological testing among suspected CD patients [49].

The multifactorial etiology of Celiac Disease

Although the exact mechanism leading to CD development is not fully understood, existing evidence suggests that CD etiology is multifactorial with genetics and gluten being the prerequisites [4, 28].

Genetics

The major genetic factors involved in the development of CD are the human leukocyte antigen (HLA) genes that encode the major histocompatibility complex class II antigen-presenting molecules on antigen-presenting cells. Approximately 95% of people with CD carry HLA-DQ2, and most of the remaining 5% express HLA-DQ8 [50, 51].
Gluten

Gluten is a water-soluble protein fraction found in wheat, barley, and rye that is highly resistant to digestion leaving some toxic fractions to penetrate the intestinal mucosa and induce an immunological reaction that results in inflammation and subsequent villous atrophy [52, 53].

Environmental and lifestyle factors

Variations in CD frequency between countries and within a given country support the notion that environmental factors contribute to CD etiology [5, 23-30, 32, 40]. Aside from gluten, other environmental and lifestyle factors such as infant feeding, socioeconomic factors, infections, antibiotics, and caesarean delivery have been proposed as summarized below [54-59].

Socioeconomic factors

Socioeconomic conditions have been shown to influence the risk of developing CD even though current evidence is contradictory. Some studies have reported that high socioeconomic conditions are associated with increased CD risk [37, 60-65]. In other studies, increased CD risk was reported in those with low socioeconomic conditions [41, 54, 66, 67], and several studies have reported no evidence for any association at all [68-71].

The exact link between socioeconomic conditions and CD development is not well established. Proposed hypotheses to explain the association between high socioeconomic status and CD include the hygiene hypothesis [72] and differences in the gut microbial [73]. The effect of socioeconomic conditions on CD development might also be pointing to other underlying lifestyle-related factors at play such as differences in the frequency of intestinal infections and a variety of other dietary factors [74, 75]. It is important to note that socioeconomic status has also been linked to the incidence of other autoimmune diseases such as Crohn’s disease and type 1-diabetes, both of which are CD comorbidities [76].

Early life events

Early life events such as caesarean delivery and preterm birth have been associated with the risk of developing CD [54, 56, 77]. During the gestational period, the fetus might be exposed to maternal risk factors. This exposure continues even during birth when newborn babies are exposed to microbes from the maternal cervix, the birth canal, and the surrounding environment, including fecal microflora [73]. The development and function of the human immune system is dependent on interactions with the human microbiome
and the mode of delivery might affect the intestinal colonization, which in turn might influence the intestinal immune responses [80, 82, 83]. There is consistent evidence showing an association between the intestinal microbiota composition and CD [73]. Interestingly, alterations in the intestinal microbiome in CD patients have been reported [84].

**Seasonal variation**

Variation in season of birth has been associated with the occurrence of CD [47, 85-87]. Two Swedish studies found an increased CD risk in children born during summer when compared to their winter counterparts [47, 85]. Tanpowpong et al. observed an increased risk of CD in children born during spring compared to any other season, with an excess risk found in boys with a spring birth [86]. This association is consistent with the theoretical model that incorporates potential environmental factors like levels of vitamin D, ultraviolet-B exposure, gluten introduction, and infections [88, 89].

**Infant feeding**

Observational studies have suggested the importance of early infant feeding practices on the subsequent development of CD [90-92]. The duration of exclusive breastfeeding and breastfeeding at gluten introduction have been associated with reduced CD risk [67, 93-95]. A window period of 4 to 6 months of age at gluten introduction has been proposed to reduce the risk of CD [91]. Sweden’s unprecedented CD epidemic was originally associated with changes to the national infant feeding guidelines both at the beginning and at the end of the epidemic [45, 93, 96, 97]. However, recent studies including two new large randomized controlled trials have contested this association by showing that the age at gluten introduction is not associated with CD risk during childhood [70, 98-100]. It has also been reported that CD is not associated with breastfeeding (any or exclusive breastfeeding) at the time of gluten introduction or with delayed gluten introduction [98, 99, 101]. It must be noted, however, that the effect of the amount of gluten at introduction has so far not been studied.

**Vaccination**

Vaccination is a plausible environmental risk factor in autoimmune disease etiology [102]. About a century ago, children received 1 vaccine for smallpox, this increased to 5 vaccines forty years ago, and today 11 vaccines are the standard routine for children by 2 years of age [103]. While this has resulted in decreased incidence of vaccine-preventable diseases, it has also raised concerns about the effect this could have on the infant’s immune system [102].
It has been hypothesized that the immunological response to vaccination among genetically predisposed individuals might lead to an immune reaction towards gluten proteins and thereby influence the risk of developing CD [102-106]. Repeated injections during vaccination might trigger an immune reaction through the stimulation of the immune system [102-106]. However, aside from the negative association between CD and tuberculosis (bacillus Calmette-Guerin; BCG), no other association has been shown in regard to vaccination [107].

**Infections**

An association between CD and lifestyle factors related to infectious load has been suggested, although the evidence for this is still contradictory. These factors include episodes of infections, the number of siblings, child daycare attendance, and socioeconomic status [41, 57, 58, 64]. Gastroenteritis has been linked to the risk of CD [108], and an American study reported an association between childhood CD and repeated rotavirus infection [58]. Other studies have also reported similar observations between CD and viral infections such as rotavirus and adenovirus [109-111]. It should be noted, however, that some studies have not confirmed this association [56, 70].

**Antibiotic use**

In the Western society, about half of all children under 15 years of age are exposed to antibiotic use at least once a year [112]. Antibiotic use affects the ecological balance of the microbiota and thus might influence the mucosal barrier, the development of the immune system, and oral tolerance [83]. Moreover, an association between intestinal dysbiosis and CD has been reported, with some studies suggesting a probable association between microbiota and CD pathogenesis. Differences in the gut microbiota between individuals with CD and healthy controls have also been reported [73]. Antibiotic use in infancy is associated with increased risk of CD development during childhood [113].
Aims of the thesis

Main Objective
The general aim of this thesis was to estimate CD incidence in Sweden and to investigate environmental and lifestyle factors that might influence the risk of developing CD during childhood because these might guide future approaches to CD prevention.

Specific Objectives
To examine the annual incidence rate of biopsy-proven CD among children in Sweden during the follow-up period from 1973 to 2009, to assess CD variation by age, sex, and birth cohorts, and to evaluate its clinical implication. (Paper I)

To investigate CD clustering at different geographical levels and to examine the association between neighborhood demographics and socioeconomic conditions and CD. (Paper II)

To explore how conditions related to maternity, delivery, and the neonatal period influence CD onset during childhood. (Paper III)

To examine whether CD is associated with season of birth or geographical place of birth. (Paper IV)
Conceptual frameworks

Several causal models have been developed to explain different causal mechanisms of diseases. This thesis is interested in assessing risk factors that are related to the development of CD, and it uses the “Life course perspective of chronic disease epidemiology framework” and the “Sufficient-component cause model” to discuss and summarize the study’s findings.

Life course perspective to chronic disease epidemiology

This framework provides an understanding of how events during different stages of life contribute to the development of health problems [114]. The life course perspective to chronic disease epidemiology is defined as the study of long-term consequences of chronic disease risk of exposures during gestation, childhood, adolescence, young adulthood, and later adult life [114]. It entails studying pathways that operate across an individual’s life course, as well as across generations, to influence the development of chronic diseases [115]. In general, the life course perspective considers two broad models of pathways between exposure and disease; the “critical period model” and the “accumulation model”. The critical period model (fetal programming) suggests that exposures act during a specific period and have permanent effects on the structure and/or functions of organs, tissues, and body systems, and these effects are not modified in any dramatic way later in life [114, 115]. In contrast to this, the cumulative model recognizes the importance of modifiers that occur later in life. It suggests that factors leading to increased disease risk or to good health accumulate gradually over the life course. There might be developmental periods where the effectiveness of these factors might have greater impact on later health than factors operating at other times. Thus, as the number and/or duration of exposures increase, there is an increase in the cumulative damage to biological systems. Environmental or behavioral insults might cause long term, gradual damage to health in separate and independent ways or they might cluster together in socially patterned ways. Such associations might reflect long-term adverse socioeconomic circumstances, maternal health and other lifestyle factors [114, 115].

The sufficient-component cause model

This model is also known as the pie-chart model and was originally proposed by Rothman. It recognizes that disease outcomes have multiple contributing determinants that might act together to produce a given disease [116]. The set of determinants that produce a specific disease in one individual might not be the exact same ones to set off the same disease in other people. For any disease
to occur, there must be a complete causal mechanism. This is also known as the sufficient cause, it is not a single factor but rather a set of factors that, if present in a given individual, will lead to the inevitable development of the disease in the individual. Rothman refers to these sets of factors as component causes, and these tend to be many for any given disease. Components of a sufficient cause do not need to occur simultaneously, and they can act at different times, thus each component cause might vary between exposure and disease development [116]. Figure 1 applies the sufficient-component cause model to describe CD causation.

**Figure 1.** Rothman’s sufficient-component cause (pie-chart) conceptual framework for discussing causality. The gray color represents necessary causes (genetic susceptibility and dietary gluten), and these are present in each of the three circles (sufficient causes) because without them CD cannot develop. The white color represents other component causes (i.e. infant feeding, infections and socioeconomic factors) that contribute to the development of CD but are not necessary for CD to occur, and these might differ from one individual to another. The question marks represent other components that are currently unknown.
**Methods**

**Overview of the study designs**

The summary of how the four papers in this thesis were designed is provided below (Table 1).

**Table 1. Overview of the study designs.**

<table>
<thead>
<tr>
<th>Description</th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
<th>Paper IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subjects</strong></td>
<td>CD cases (9 107)</td>
<td>CD cases (2 080)</td>
<td>CD cases (6 569)</td>
<td>CD cases (6 569)</td>
</tr>
<tr>
<td></td>
<td>Children born (1 578 094)</td>
<td>SAMS areas (8 036)</td>
<td>Total population (1 912 204)</td>
<td>NUTS 1 regions (3)</td>
</tr>
<tr>
<td></td>
<td>Person-years (28 039 741)</td>
<td>Hospitals (43)</td>
<td>NUTS 2 regions (8)</td>
<td>Person-years (7 768 096)</td>
</tr>
<tr>
<td><strong>Study design</strong></td>
<td>Prospective longitudinal</td>
<td>Ecological</td>
<td>Prospective longitudinal</td>
<td>Prospective longitudinal</td>
</tr>
<tr>
<td><strong>Registers</strong></td>
<td>The National Swedish Childhood CD Register</td>
<td>The National Swedish Childhood CD Register</td>
<td>The National Swedish Childhood CD Register</td>
<td>The National Swedish Childhood CD Register</td>
</tr>
<tr>
<td></td>
<td>Statistics Sweden</td>
<td>ASTRID database</td>
<td>Medical birth register</td>
<td>Medical birth register</td>
</tr>
</tbody>
</table>

CD: celiac disease; NUTS: Nomenclature of territorial units for statistics; SAMS: Small area market statistics

**Study setting**

Sweden has a population of 9.7 million inhabitants with about 100 000 births annually [117]. Sweden’s maternity and child health care system is decentralized, free, and actively offered to all expectant mothers, newborns, and pre-school children [118]. Information from the health care system and other administrative units is stored in national registers that are under the administration of several different government agencies. This information is
available for research purposes after obtaining ethical approval. Linking several registers is possible and is conducted by Statistics Sweden using a personal identification number (PIN). Every individual residing in Sweden for \( \geq 1 \) year is assigned a PIN, which consists of twelve digits, the first eight representing the date of birth (year-month-day) followed by three digits that identify the individual and the last digit that is for control purposes.

**Geographical area classifications**

Investigations in this thesis were carried out on varying geographic levels. All papers took into consideration the national level. Two levels of the European Union (EU) nomenclature for territorial units of statistics (NUTS 1 and NUTS 2 regions) were explored in Papers IV and II, respectively. NUTS is a geocode standard developed and regulated by the EU for referencing subdivisions of countries for statistical purposes. The NUTS 1 regional level divides Sweden into three regions while the NUTS 2 regional level divides the country into eight regions. Other levels considered in Paper II are hospital catchment areas, municipalities, and the Small Area Statistics (SAMS areas). SAMS areas were used as the basic unit of measurement and are viewed as statistical units. The formation of SAMS areas was mainly based on municipal planning zones and voting districts representing homogenous residential neighborhoods in terms of demographic and socio-economic characteristics. This thesis uses the term “neighborhood” synonymously with SAMS area, and these are often used interchangeably.

**Data source**

Data for this thesis were obtained from the Swedish Initiative for Research on Microdata in the Social and Medical Sciences (Umeå SIMSAM Lab), which is part of the national Swedish Initiative in the Social and Medical Sciences [119]. This research infrastructure links microdata obtained from several national registers administered by Statistics Sweden and the National Board of Health and Welfare. The Umeå SIMSAM Lab seeks to perform high-quality interdisciplinary microdata research on the interplay between childhood and societal aspects of lifelong health and welfare. The Umeå SIMSAM Lab hosts an exceptional microdata infrastructure with longitudinal data on demographics, socioeconomics, and health indicators for the entire Swedish population covering several decades. The Umeå SIMSAM Lab is hosted by Umeå University and is coordinated through several Umeå University departments. In this thesis, Papers III and IV used data obtained directly from the Umeå SIMSAM lab. For paper I, data were obtained from the National Swedish Childhood CD Register, and data for paper II were obtained from the
ASTRID database, both of which are hosted by Umeå University and are currently part of the Umeå SIMSAM Lab. Below is a brief description of the registers from which data for this thesis were obtained.

**The National Swedish Childhood CD Register**

Identification of CD cases included in this thesis was based on recruitment of all children diagnosed with CD aged \( \geq 14.9 \) years and reported to the National Swedish Childhood CD Register. This register was established in 1991 and is hosted by the unit of Epidemiology and Global Health at Umeå University, Sweden. The establishment of this register came out of the need to track changes in the number of children diagnosed with CD throughout Sweden prompted by an abrupt rise in the incidence of CD during the mid-1980s [45]. Children diagnosed with CD between 1973 and 1990 were retrospectively added to the register from the records of five pediatric units, which was equivalent to 15% national coverage. From 1991 to 1997, 14 pediatric units (40% national coverage) reported their cases to the register. In 1998, nationwide coverage was attained, and all 47 pediatric units reported their CD cases to the register [46].

The CD status of children reported to the National Swedish Childhood CD Register was determined using the diagnostic criteria provided by ESPGHAN. For cases identified before 1990, the diagnostic criteria were established in 1970 and required three consecutive small intestinal biopsies first with a normal gluten-containing diet, second without gluten, and third after re-introduction of gluten [120]. The new diagnostic guidelines were launched in 1990 by ESPGHAN and requires that an individual shows villous atrophy while on a normal diet followed by clinical remission on a gluten-free diet. This criterion were not fully adopted in Sweden until about 10 years later.[121]

**ASTRID database**

The ASTRID database is a longitudinal and geo-referenced individual database containing demographic and socioeconomic data on the entire Swedish population and was established in 1985. Information in this database is obtained from several registers hosted by Statistics Sweden. The database contains information including details of workplaces and homes that provides the opportunity to combine environmental data on a personal level with different variables on regional conditions, local conditions, neighborhoods, homes, and workplaces. Data from this register were used in Paper II.
The Swedish Medical Birth Register

The Swedish Medical Birth Register was established in 1973 by an act of the Swedish Parliament, and it is maintained by the Swedish National Board of Health and Welfare [118]. The main purpose of this register is to compile information on antenatal and perinatal factors to help assess how these factors influence the health of infants. The register contains ~98% of prospectively recorded data on all newborns in Sweden [122]. The content and methods of data collection have evolved over time, although the basic structure and content have remained the same. Overall, it contains three records of primary interest; the basic antenatal care record of the mother, the delivery record, and the record for the pediatric examination of the newborn infant [122]. According to the National Board of Health and Welfare, only 1–2% of the records are missing for most of the years, and this is considered acceptable [123]. Data from this register were used in Papers III and IV.

The LISA database

The Longitudinal Integration Database for Health Insurance and Labor Market Studies (LISA database) was established in 1990. The database includes all individuals aged 16 years and older who are registered residents of Sweden by December 31 of each year [117]. The register uses background information obtained from population registers, and it also gathers information from other registers on labor market, education, and social sectors. This database primarily focuses on individuals, but it also includes information on family connections, employers, companies, and the location of the workplace. Data from this register were used in Paper III.

Study variables and statistical analysis

Paper I

In this cohort study, the main variables of interest were year of birth and age at diagnosis. CD frequency was estimated as incidence rate, cumulative incidence, and clinical impact. The annual incidence rate was calculated by dividing the number of new cases diagnosed by the number of person-years of follow-up approximated by the mid-year population. The cumulative incidence of each birth cohort at a specified age was estimated by dividing the number of cases diagnosed up to this age by the total number of births in the cohort. Incidence rates are presented as the number of new cases per 100 000 person-years, and the cumulative incidence is reported as the number of cases per 1 000 births. The clinical impact was defined as the annual number of new CD cases diagnosed throughout the country and was reported by age groups.
The statistical significance of differences between incidence rates was determined through logistic regression. The log-rank test was used to test differences in statistical significance of the cumulative incidence between birth cohorts. Statistical significance was defined as a \( p \)-value of <0.05. Analysis was performed using Microsoft Excel 2008 (version 12.3.2) and SPSS 19.0 for Windows (SPSS, Chicago, IL).

**Paper II**

This study applied a multilevel approach, and four levels were considered:

i) the NUTS 2 regions, ii) hospital catchment areas, iii) municipalities, and iv) SAMS areas. We included 8 NUTS 2 regions, 43 hospital catchment areas, 253 municipalities, and 8 036 SAMS areas. SAMS areas were the statistical units used in the analysis.

Demographic variables included population density, which was generated by dividing the number of inhabitants per SAMS area unit by its area expressed in square kilometers. Age was estimated as the population’s summed age in years divided by the number of persons in each SAMS area. Region of origin was calculated by first summing up all individuals born in the Nordic and non-Nordic countries, respectively, and dividing them by the total population. Family size was approximated to the average number of children per female aged 18 years or older.

Socioeconomic variables were education, employment, income, and level of economic activity in the area of residence. The highest and lowest level of education was considered for residents aged >20 years. Compulsory education was defined as the percentage of residents with 9 years of schooling as the maximum education level attained. University education was defined as the percentage of residents that completed >2 years of university education. Percentage employed was derived by computing the total number of persons aged 16–64 years who were employed divided by the total population in that age range. Average income was the mean income from wages and self-employment for residents aged 16–64 years expressed in 100s of SEK. The level of industrial and commercial activity was derived by calculating the ratio between the total number of persons working within a neighborhood (the numerator) and the total sum of this value and the additional residential population (the denominator).

Multilevel logistic regression analyses were performed, and the binary response variable was constructed by taking the mean SAMS incidence rate as the cut-off. Results from the fixed parts of the models were presented as odds ratios (ORs) with 95% confidence intervals (CIs), and statistical significance was defined as a \( p \)-value <0.05. For random effects, comparisons
were between the intraclass correlation coefficient and the median odds ratio. All analyses was performed using Microsoft Excel 2008 and STATA version 11.2 software.

**Paper III**

Several maternal characteristics during pregnancy were included in the analysis. Age was grouped into <25, 25–29, 30–34, and ≥35 years. Disposable income was categorized in three strata based on quartile ranges. Smoking was categorized as yes or no. Body mass index (BMI) measured at the first antenatal visit was divided into underweight (<18.50), normal weight (18.50–24.99), overweight (25.00–29.99), and obesity (≥30.00). Parity was 1 if the index child was the first and the mother had never given birth, 2 if the current child was the second child, and 3 if the index child was the third child or more. Duration of pregnancy was either gestational age <37 weeks or ≥37 weeks. Mode of delivery was categorized into caesarean or vaginal, and caesarean delivery was further divided into elective or emergency. Maternal infections and premature rupture of the membrane (PROM) were defined according to the International Classification of Diseases (ICD) 9th and/or 10th revisions and were categorized as yes or no [124].

Several infant characteristics were analyzed. Birth weight in grams was categorized as very low birth weight (<1 500 g), low birth weight (1 500–2 499 g), or normal birth weight (≥2 500 g). Sex was either male or female. Small for gestational age was categorized as yes or no. Apgar score at five minutes after birth was recorded as either low (<7) or normal (≥7). Infant neonatal infections were classified according to the ICD 9th and/or 10th revisions [124] and categorized as yes, if any infection was reported and no, if there were no infections reported.

Analysis was performed by using two-by-two tables and logistic regression, and results were reported as ORs with 95% CIs. Statistical significance was defined as a *p*-value of <0.05. All statistical analyses were performed using SPSS 22 for Windows.

**Paper IV**

Season of birth was divided into winter (December, January, February), spring (March, April, May), summer (June, July, August), and autumn (September, October, November). Geographical location was classified according to the Swedish NUTS 1 regions, which include northern, central, and southern Sweden. Age at diagnosis was defined as the age when the first positive small intestinal biopsy was obtained and was categorized as 0–1.9 or 2–14.9 years. Year of birth was divided into three cohorts generated based on our previous knowledge on differences in CD risk: the birth cohort of 1991–
1996 represented the epidemic years, the 1997–2002 birth cohort represented children born immediately after the epidemic, and the 2003–2009 cohort represented children born when the CD epidemic had completely abated.

The Cox proportional hazards regression model was used to determine associations between the independent variables and age at diagnosis. Results were reported as hazard ratios (HRs) with 95% CIs. Statistical significance was defined as a $p$-value of $<0.05$. Analyses were performed using SPSS 22 for Windows.

**Ethical considerations**

The Regional Ethical Vetting Board in Umeå approved all research based on data from the Umeå SIMSAM Lab, including the present project.
Results

The main findings in this thesis are presented in three sections. The first section presents CD incidence rates as summarized from Papers I. The second section describes the geographical variation in CD from Papers III and IV. The third section presents the findings on CD risk factors examined at the neighborhood and individual levels (Papers II, III, and IV).

Incidence rate

Annual overall incidence rates for 1973 to 2009

Over the follow-up period, the average annual incidence rate increased by 4% in the child population except for a temporal decline from 1995 to 1997 (OR = 1.04, \( p < 0.001 \)) as shown in Figure 2 (Paper I).

Figure 2. Age-specific annual incidence rates from 1973 to 2009.
**Age-specific annual incidence rate**

Pronounced variation in incidence rate was observed in children diagnosed before the age of 2 years, with an epidemic pattern evident in 1984–1995. However, at the end of the follow-up period this group had the lowest incidence. A gradual but persistent increase in the incidence was noted in the age groups of 2–4.9 years and 5–14.9 years (Figure 3) (Paper I).

![Age-specific annual incidence rates](image)

**Figure 3.** Age-specific annual incidence rates according to the age groups of 0–1.9 years, 2–4.9 years, and 5–14.9 years.

**Age at diagnosis**

The median age at diagnosis was initially stable at about 1.1 years, but by 1995 it had increased to 4.6 years and at the end of the follow-up it was about 6.7 years.
Cumulative incidence

The epidemic birth cohorts (1984–1994) maintained the highest cumulative incidence throughout the entire follow-up period except for the 2000–2002 birth cohorts (Figure 4). The cumulative incidence of other birth cohorts gradually increased as the birth cohorts grew older this is shown in Figure 4 (Paper I).

Figure 4. The cumulative CD incidence for children born between 1973 and 2009. The cumulative incidence of 1973–1997 is aggregated into three intuitive groups according to similarities in the cumulative incidence. From 1998 to 2008, the cumulative incidence is reported separately for each year.

Geographical variation in Celiac Disease

NUTS 1 region level

On the NUTS 1 regional level, the risk of developing CD was increased for children born in the southern (adjusted HR = 1.59; 95% CI = 1.48–1.72) and central (adjusted HR = 1.19; 95% CI = 1.10–1.29) regions compared to children born in the northern region (Paper IV).
Seasonal variation in incidence by NUTS 1 regions

The pattern in season of birth differed according to region. In southern Sweden, the CD incidence rate began to increase from December to May except for a slight drop during April. For the remaining months of June to November, the incidence rate appeared to drop, aside from a slight increase in August. In central Sweden, from October to January the incidence rate was low, and an increase started from February to May, dropped again in June, but increased in August and September. In northern Sweden, the CD incidence rate remained low for most of the months except for an increase in April, May, and July (Paper IV).

Figure 5. Incidence rate by month of birth for the three NUTS 1 regions of Sweden.
**SAMS incidence rate**

The mean SAMS-area CD incidence rate was 28 cases/100 000 person-years. The incidence results are presented in Figure 6, which maps the SAMS-level CD incidence rate. Figure 7 provides detailed examples of the spatial distribution of CD incidence at the NUTS 2 region level (South Sweden) with a snapshot of Malmö municipality (Paper II).

![Figure 6. SAMS-level CD incidence rates (<15 years of age) 1998–2003. Cases per 100 000 person-years.](image_url)
Figure 7. SAMS-level CD incidence rate in Malmö in southern Sweden. The color scheme in this figure is similar as that in Figure 6.
Environmental factors associated with Celiac Disease

Table 2 summarizes the environmental/lifestyle factors investigated in this thesis with regard to their association with CD development. These factors are presented with respect to the level at which the association was investigated, for example, SAMS areas (neighborhoods) and the individual level. The table provides information on exposures investigated for each paper.
## Table 2. Summary of the associations between environmental factors and CD as determined in this thesis.

<table>
<thead>
<tr>
<th>Level</th>
<th>Exposure</th>
<th>Association with CD</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neighborhood</td>
<td>Population density</td>
<td>No association</td>
<td>Paper II</td>
</tr>
<tr>
<td>Age</td>
<td>Lower risk with higher age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country of origin</td>
<td>No association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family size</td>
<td>No association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>Lower risk with higher education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td>No association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work income</td>
<td>Higher risk with higher income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrial &amp; commercial areas</td>
<td>Lower risk in industrial and commercial areas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual</td>
<td><strong>Maternal characteristics</strong></td>
<td></td>
<td>Paper III</td>
</tr>
<tr>
<td>Age</td>
<td>Lower risk with age ≥35 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disposable income</td>
<td>Lower risk with higher income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking in pregnancy</td>
<td>No association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>Lower risk in boys with high BMI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infections</td>
<td>Higher risk in girls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of pregnancy</td>
<td>No association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROM</td>
<td>Lower risk in boys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elective caesarean section</td>
<td>Higher risk in boys</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Infant characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity</td>
<td>Higher risk for second born</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight</td>
<td>Lower risk in boys &lt;1 500 g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apgar score at 5 minutes</td>
<td>Lower risk in girls with &lt;7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neonatal infections</td>
<td>No association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season of birth</td>
<td>Lower risk in winter</td>
<td></td>
<td>Paper IV</td>
</tr>
</tbody>
</table>

BMI: body mass index; CD: celiac disease; PROM: premature rupture of the membrane
Discussion

Main findings

The CD incidence in children aged 0–14.9 years gradually increased throughout the study period, aside from a temporary decline in 1996–1997 (Paper I). The median age at diagnosis increased from 1.0 to 6.8 years by the end of the follow-up (Paper I). The highest risk of developing CD was found in children born in southern Sweden, followed by those born in central Sweden (Paper IV). Variation in CD incidence was more pronounced at the municipality level compared to the NUTS 2 regional level (Paper II). On the SAMS level, a positive association was found between work income and CD risk, and the risk was significantly reduced in areas with increased average age and areas with a high level of industrial and commercial activity (Paper II). At an individual level, and regardless of sex, a reduction in CD risk was observed among children with mothers aged ≥35 years and among children with mothers earning high incomes (Paper III). Among boys, elective caesarean delivery increased CD risk, while maternal overweight, PROM, and low birth weight showed a negative association. Among girls, the risk of developing CD was increased by repeated maternal urinary tract infections during pregnancy (Paper III). For season of birth, spring birth was associated with increased CD risk in children diagnosed at <2 years and in the birth cohort of 1991–1996, while the risk for CD was higher in those diagnosed at 2–14.9 years if they were born during summer or autumn compared to winter births (Paper IV).

Strengths and limitations

Data for this thesis were obtained from Swedish national registers of high quality and with access to data for all children in the country. Cases were identified prospectively through the National Swedish Childhood CD Register. We are certain that variation in hospitals reporting to the register did not significantly affect the reported incidences because our incidence findings concur with an earlier study that had full national coverage [39]. Moreover, we also previously reported that 40% coverage was sufficient to estimate the national CD incidence [46]. Because this thesis had a large number of cases and covered the entire child population, it was possible to adjust for several factors. One drawback, however, is that for Papers II, III, and IV we excluded cases that were reported without a PIN because such cases could not be linked to other registers. Nevertheless, even after this exclusion the remaining population was large enough to provide strong evidence in
favor of the observed associations. Misclassification of cases as non-cases might have occurred because children without a PIN were coded as non-cases, which might imply underestimation of the incidence and risk estimates.

CD diagnosis was ascertained by villous atrophy and cases without villous atrophy were excluded irrespective of whether they had elevated serological markers, minor enteropathy (Marsh I), and/or clinical symptoms typical for CD. Besides, the latest ESPGHAN guidelines on serological markers for CD diagnosis were introduced in 2012, which was long after this follow-up [125]. In fact, the strict diagnostic criteria applied might imply that CD incidence was underestimated in this thesis. Moreover, the Swedish National Childhood CD register primarily captures cases clinically detected, yet about 2/3 of all cases remain undiagnosed [48]. It is possible, therefore, that there are differences between clinical and subclinical CD cases.

One general drawback of using register data in research is that this information is primarily collected for other governmental purposes, thus there is no room for including other variables of interest if they were not originally collected. In our case, it would have been interesting to include data on infant feeding. However, this was not possible because this information was not available. On the positive side, however, because data on CD and on risk factors were collected separately before this study was conceived, our findings do not suffer from certain biases, e.g. recall bias or interviewer bias, that are often associated with other data-collection approaches.

Our investigations on the neighborhood level (Paper II) employed an ecological approach, which is often regarded as the lowest degree of evidence in observational studies. However, at the time this study was conducted this was the only source of data available. Besides, the role of neighborhood conditions in CD etiology have not been assessed before even though such factors have been shown to affect health [126]. It is also true that an ecological study design is prone to the “ecological fallacy” where conclusions about individuals are based on associations established at an aggregated level. We recommend that these finding be interpreted at an area level to avoid this problem. It should also be noted that area-based measures are valid depending on whether the geographical unit is meaningful rather than whether or not it serves as an adequate proxy for individual-level measures [126]. In this study, SAMS areas are a varied geographical unit not only in Sweden but also in other Nordic countries.
Celiac Disease incidence is increasing

We found an overall increase in CD incidence mainly due to the rising incidence in children diagnosed between the ages of 2 to 14.9 years. This observation is in line with current evidence showing an increase in CD incidence [71, 127, 128]. Increased awareness is one of the contributing factors to this increase, but it is not the sole explanation [128, 129]. The increase in serological testing during general routine screening might have contributed to the observed increase in CD.

The median age at diagnosis has increased throughout the follow-up period. The upward shift in age at diagnosis can partly be explained by growth of the epidemic birth cohort of 1985–1995, and as they grow older they carry with them an increased risk. Similar observations have also been shown elsewhere and may perhaps indicate that CD is gradually changing from presenting in early childhood to later age at presentation [130]. There must be other factors associated with this increase, however, because similar patterns are also observed in post-epidemic birth cohorts. Prolonged exposure to environmental factors might contribute to disease progression in older children and adults carrying low-risk HLA gene types with a lower genetic burden who are known to seldom develop CD. Continuing increased awareness of CD symptoms typical in older children by medical personnel and caretakers, as well as screening of siblings, could also have contributed to the shift towards older age at diagnosis. However, the role of additional environmental exposures should not be overlooked.

Geographical variation in Celiac Disease risk

We assessed CD risk between the three NUTS 1 regions, and children born in southern Sweden had the highest CD risk, followed by those living in central Sweden. The Swedish population is relatively homogenous with a high prevalence of HLA risk genes, and about approximately 50% of the population carries such genes [51]. Vitamin D deficiency is suggested to explain seasonal variation in CD risk [88]. Seasons tend to vary geographically, and it is possible that the geographical variation in CD risk is related to seasonal differences. Geographically, Sweden stretches about 1 570 km from south to north, and it lies within latitudes 55°–69°, which affects its exposure to sunlight. Thus people living in northern Sweden receive the fewest hours of sunlight exposure especially during the winter [131], yet they have the lowest incidence of CD. Another new and controversial hypothesis suggests a link between high vitamin D levels and increased CD risk [89]. This theory is based on comparisons between countries, but no data on individuals have been
analyzed to ascertain this association [89]. Additionally, optimal vitamin D levels were shown for both the child and adult population in Sweden [131, 132]. Thus, perhaps the Swedish north to south gradient and the difference in CD by season of birth can be explained by other environmental factors such as infections and infant feeding habits.

Further investigations revealed that CD variation was more pronounced at the municipality level compared to NUTS 2 regions. This is not surprising given the fact that municipalities are closer to individuals compared to NUTS 2 regions and therefore might better reflect homogeneity of physical living conditions and policy implementations such as daycare attendance and infant feeding recommendations. Other factors that might possibly explain this variation are differences in eating habits, differences in the composition of microbiota, and variations in the identification of CD patients by local health care facilities. Also, we cannot rule out differences in health-seeking behavior.

The life course perspective and Celiac Disease risk factors

To date, we are sure of two factors that are necessary for the development of CD – genetic susceptibility and gluten exposure. By employing the life course model for chronic disease epidemiology, the findings from this thesis can be discussed with a goal of highlighting how this thesis contributes to our knowledge of CD etiology. Five life stages are specified (fetal life, delivery, infancy, childhood, adulthood) with death as the end point (Figure 8). Factors investigated in this thesis contribute to the first two stages and will be discussed in detail below.
Figure 8. A schematic representation of biological and environmental exposures acting across the life course that might influence the development of CD. SES: Socioeconomic status; PROM: premature rupture of the membrane.
Fetal conditions and Celiac Disease

Socioeconomic conditions during pregnancy

Socioeconomic conditions have been a factor of interest in relation to CD, although existing evidence is contradictory as exemplified in Table 3. This thesis found high average neighborhood income to be associated with an increased risk of developing CD, while high maternal income was associated with reduced risk of CD in the children. We are not certain of the reasons why the effect of income differed at the neighborhood and at the individual level, although similar contradictions have been reported before [37, 60-65]. One possible explanation is that the hygiene hypothesis used to explain the association between socioeconomic conditions and CD fits better when making comparisons on area levels, i.e. between neighborhoods and between countries, than at an individual level. Interestingly, some studies that used socioeconomic data on the area level such as Kandroshova et al., Zingone et al., and Whyte et al. also reported higher risk of CD in children belonging to higher socioeconomic groups, similar to the observation from our ecological study (Paper II) [37, 65, 66]. It is also possible that the observed effect of neighborhood-level income mirrors differences in other lifestyle factors such as eating habits. Another possible explanation is that high-income earners use medical care more often, which would translate into more CD diagnoses. However, this is not very likely because access to pediatric care is free for all children <18 years in Sweden. [133].

The observed reduced risk associated with high maternal income and neighborhood education might be a proxy for environmental exposures such as infections and breastfeeding [134]. The hygiene hypothesis suggests that children with poor socioeconomic background are more likely to be exposed to infections earlier in life that help to challenge the immune system [37]. Early exposure to infections challenges the immune system to develop, which later prevents autoimmunity and thus reduces the risk of developing CD [72, 104]. However this hypothesis is not backed by recent evidence on the association between CD and early childhood infections [70, 107], nor by our observations of CD and maternal income (Paper III) or several of the other studies indicated in Table 3 [41, 54, 66, 67].

The conflicting findings on socioeconomic conditions and CD development might be due to methodological differences in study design, definition of exposure variables, study population, and setting. For example, in our case “work income” does not include other sources of income, thus individuals with significant personal wealth might be classified as low-income earners. Additionally, it is possible that each measure of socioeconomic conditions
yields a different effect. Based on our findings on income and on the existing literature, it is not possible to draw conclusions about the role of socioeconomic conditions in CD etiology.
Table 3. Literature review on the association between socioeconomic conditions and celiac disease risk.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Country</th>
<th>Age group</th>
<th>Study duration</th>
<th>Source of data on socioeconomics (SES)</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased CD risk due to high socioeconomic status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Namatovu et al, 2014</td>
<td>Sweden</td>
<td>0-15</td>
<td>1998-2003</td>
<td>Average area; work income, occupation and education level</td>
<td>CD high in high annual income areas Education reduced the risk</td>
</tr>
<tr>
<td>West et al, 2014 [83]</td>
<td>UK</td>
<td>0-70+</td>
<td>1990-2011</td>
<td>Indices of multiple deprivation</td>
<td>Lower CD in most deprived area</td>
</tr>
<tr>
<td>Whyte et al, 2014 [78]</td>
<td>UK</td>
<td>0-16</td>
<td>1995-2012</td>
<td>Welsh multiple deprivation index</td>
<td>CD prevalence highest in the highest SES population</td>
</tr>
<tr>
<td>Olen et al, 2012 [86]</td>
<td>Sweden</td>
<td>16-64</td>
<td>1999-2008</td>
<td>European SES classification based on occupation</td>
<td>Lowest social class, less diagnosed with CD</td>
</tr>
<tr>
<td>Kondroshova et al, 2008 [43]</td>
<td>Finland and Russia</td>
<td>6-18</td>
<td>Finland:1994; Russia: 1997-2001</td>
<td>Socioeconomic conditions</td>
<td>High CD in wealthier Finland than Russian Karelia</td>
</tr>
<tr>
<td>Fera et al, 2003 [85]</td>
<td>Italy</td>
<td>20-70</td>
<td>2000-2001</td>
<td>Education and employment status</td>
<td>CD higher in employed versus housewives, student and retired</td>
</tr>
<tr>
<td>West et al, 2003 [84]</td>
<td>UK</td>
<td>45-76</td>
<td>1990-1995</td>
<td>Occupations</td>
<td>CD low in partly skilled/unskilled workers compared to professionals</td>
</tr>
<tr>
<td>Reduce in CD risk due to socioeconomic conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No association</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White et al, 2013 [81]</td>
<td>UK</td>
<td>0-16</td>
<td>1990-2009</td>
<td>Multiple deprivation index</td>
<td>Similar scores in CD cases and general population.</td>
</tr>
<tr>
<td>Welander et al, 2010 [77]</td>
<td>Sweden</td>
<td>0-8</td>
<td>2007-2008</td>
<td>Maternal education</td>
<td>No difference between classes</td>
</tr>
</tbody>
</table>

CD: Celiac disease; SES: Socioeconomic Status; UK: United Kingdom
**Maternal infections during pregnancy**

We also found increased CD risk to be associated with maternal infections during pregnancy, although the causal mechanism remains to be identified. Further analysis clarified that this association was mainly due to repeated urinary tract infections rather than to other maternal infections. Antibiotics often used in treating urinary tract infections are suspected to influence the maternal and later the infant gut microbiota, and this might increase the risk of CD [135]. Additionally, antibiotics have also been shown to increase CD risk [55]. We further compared children with mothers aged ≥35 years with high income and no urinary tract infection versus the other mothers and found synergistic effects on the risk for CD. We found no association between neonatal infections and CD, although some earlier studies have also shown a positive association [57, 70].

**Age during pregnancy**

We reported reduced CD risk in neighborhoods with older residents and in the offspring of mothers aged ≥35 years. The reduced CD risk in neighborhoods with high age might be due to the fact that many of the children of these residents are ≥15 years and therefore were not included in our study. The reduced incidence associated with maternal age at an individual level might indicate that age influences lifestyle choices, for example, older mothers are more likely to practice exclusive and prolonged breastfeeding [136-138]. Breastfeeding and certain infant feeding practices have previously been associated with reduced CD risk [67, 90, 91, 93], although recent evidence does not support this association [98-100]. A randomized controlled trial would be suitable to clarify the effect of breastfeeding on CD, but this is impossible due to ethical reasons. It is also possible that our findings are a proxy for other factors yet to be identified; therefore, our observations are not conclusive.

**Smoking during pregnancy**

No association was found between maternal smoking and the risk of developing CD. This finding is in agreement with other Swedish studies that also investigated smoking during pregnancy and CD development in the offspring [59, 139-141]. Some studies on the adult population have even reported smoking to be protective against CD development [68, 142]. It is possible that smoking has an effect, but it appears that this effect does not play any role in CD development in the Swedish setting because even a Swedish study on the adult population found no association between CD and smoking [140].
**Body Mass Index**

Some studies have suggested that high fat dietary environment during the fetal period influences immune development which may lead to lifestyle related disease later in life [143]. In our study overweight and obesity were associated with reduced risk in boys but not in girls. Although this association was weak, thus it’s possible that this is a chance finding and warrants further investigations.

**Delivery factors associated with Celiac Disease risk**

Caesarean section was not a risk factor in this study, although a slightly increased CD risk in boys was associated with elective caesarean delivery. Another study reported a similar association, although they did not stratify for sex [56]. Why would boys born by elective caesarean section have a increased risk? Dysbiosis has been used to explain the association between CD and caesarean delivery. Children born by elective caesarean delivery are the only ones who completely evade exposure to the maternal vaginal and fecal microbiota that usually occurs during delivery. This might result in these children developing a different initial gut colonization and subsequently disturbed intestinal immune maturation and mucosal barrier function [144], which in turn might affect their risk of CD. This finding is further complemented by our finding that PROM was associated with reduced CD risk. Moreover, it’s possible that the reason a similar association is not observed in relation to emergency caesarean delivery is that ruptured membranes expose the baby to the maternal vaginal microbiota for a relatively longer duration than the baby would be even in vaginal delivery. The vaginal microbiota of pregnant women has been shown to be rich in lactobacillus species [145], and it is also possible that exposure to lactobacilli during infancy affects the early colonization of the infant’s gastrointestinal tract. Interestingly, lactobacilli have been shown to have a modulating effect on the immune response to gluten in cell culture and mouse models, and this has contributed to our speculations that PROM might be associated with reduced risk of CD [75].

Compared to being the first child, high parity was associated with increased CD risk, although the risk was only significant for second born children. This association may be due chance since parity was used as a measure for infection, with higher parity indicating exposure to higher infectious load brought to the index child by older siblings. In a previous case-referent study number of siblings was not associated with CD risk [111]. The risk of developing CD was more prominent in girls compared to boys. We report that elective caesarean delivery, maternal infections, PROM and Apgar score
appeared to affect the risk of developing CD differently for boys and girls. Although at the moment the reasons for these differences are difficult to explain. This finding might provide shed some light on reasons behind the huge difference in the risk of developing CD between boys and girl. Thus, further investigations are needed to test such a hypothesis.

**Other stages of the life course**

There are other factors that are of interest in CD development that can be identified at other stages of the life course but that we have not been able to investigate because of we lacked this data. Even among the two stages investigated in this thesis, some factors would have been interesting to include, for example, infant feeding, viral infections, maternal smoking while breastfeeding and daycare use because there is currently not enough evidence to clarify their etiological importance. There appears to be even less known about the environmental risk factors associated with CD during adult life.

**The sufficient-component cause model and Celiac Disease risk factors**

In understanding CD etiology, the sufficient-component model plays an important role because it addresses the issue of multifactorial causation that is evident in CD etiology. Genetics and the consumption of gluten fit the criteria for the necessary components of this model because they are prerequisites for CD to occur. As this model suggests, these factors are necessary but are not sufficient for CD to develop, and other contributing risk factors are required. Environmental factors such as those investigated in this thesis – including socioeconomic conditions, infections, and mode of delivery – qualify to be considered as component causes. All component factors do not need to be present for CD to occur, and different combinations of factors act together and result in CD. Therefore, the sufficient causes, i.e. a combination of both the necessary and the component causes that produce CD in one individual, might not necessarily trigger CD in another. Even though there currently appear to be several examples of component causes in the multifactorial etiology of CD, we speculate that there are even more that have yet to be identified.
Concluding remarks and future research

This thesis is in line with current evidence that suggests an increase in CD frequency as demonstrated by the continually increasing incidence in Sweden’s child population. More and more children are getting their diagnosis at a later age, which is a major shift from the majority of cases being diagnosed at <2 years of age as was the case previously. One possible explanation for this might be that children diagnosed at <2 years have a high genetic risk and so they do not need several other environmental exposures to trigger CD development. On the other hand, those diagnosed later might have a lower genetic risk and thus need high and/or repeated exposure to other risk factors for CD to occur, thus explaining the increased risk as they grow older.

This thesis has contributed new knowledge by identifying factors that contribute to the risk of CD at the neighborhood level, including income, education, and age. We have found that caesarean section and maternal UTI are associated with the risk of developing CD both findings were sex specific and are related to dysbiosis. Differences in CD risk between boys and girls have been reported before and justify separate analysis in CD studies. The well-known higher CD risk in girls seen long before puberty is a very interesting phenomena that should be further studied. We also found that maternal age and income with possible association to life style were associated with reduced CD risk. Moreover, all factors could be linked to the development of the immune system. In future, preventive strategies could benefit by taking measures to prevent gut dysbiosis. Reduced prescription of antibiotics is one way that would work against the problem with development of bacterial resistance to antibiotics. To promote normal delivery would also be in line with goals in the field of maternal and child health. To some of the other findings we cannot find a mechanistic explanation and it is worth to remember that one important role of epidemiological studies is to generate new hypothesis and thereby contribute to future search for disease mechanisms and possible treatment and prevention.

Put together, our findings illustrate that factors at all stages of the life course might influence the risk of developing CD. It is also possible that a given exposure might affect CD risk in more ways than one because of the wide range of exposures and the importance of timing for any given exposure. Factors on different societal levels such as the national, regional, neighborhood, and individual level affect the risk of developing CD. There is a need, therefore, to move beyond investigations that focus on an individual level and to incorporate multilevel investigative approaches. For CD prevention to be possible, a general approach is needed to reach as many
people as possible. We still do not know the specific time during the life course that can be considered the critical window for developing CD; therefore, researchers still need to investigate all of the different life stages. Additionally, to provide more convincing evidence for the risk factors for CD, future research would benefit from using data with repeated measures such as register-based data to test the life course theory by examining CD risk at different points in time. Having access to multilevel data, i.e. on an individual, neighborhood, regional, and national level, will provide an added advantage because this can clarify the effect of factors at all levels. Ongoing prospective cohort studies will also contribute to the understanding of the disease process.

This thesis conforms to Rothman’s pie-chart theory by showing that in the case of CD, genetic susceptibility and exposure to gluten are the necessary causes without which CD cannot occur. However, these two are not sufficient and so other environmental factors (component causes) such as the ones investigated in this thesis are important. However, from the findings of this thesis and the existing literature it is clear that we have not yet identified all of these components. This thesis differs from Rothman in the use of the term “cause”, for there is not enough supporting evidence to characterize the factors investigated in this thesis as causes, and therefore it is preferable to continue referring to them as risk factors until further clarification.

A question that has lingered in my mind during the entire process of my doctoral training is whether this thesis adds to existing knowledge on CD. This is an especially important question considering the lessons from the banana diet, which was first hailed for its ability to treat CD following its introduction by Dr. Sydney Hass in the 1920s, yet years later it was discovered that it was not the banana diet but rather the exclusion of wheat and barley that the diet required. Somewhat similar examples can be seen even in recent research, such as the current debate on the association between infant feeding practices such as breastfeeding and the incidence of CD. In the wake of the conflicting findings seen in this thesis and in the literature, it is hard to speculate on how the future will view the current evidence. Nevertheless, lessons can be drawn and knowledge on CD and research in general are obtainable.
The researcher

My first encounter with what I now understand as research happened while doing an assignment in a human rights course as an undergraduate training in bachelor's degree in Arts majoring in history. The assignment was to write a case report on any human rights issue of interest. At that time, Ugandan newspapers and media outlets where discussing the infamous case of a pedophile who had brutally defiled a six-month-old girl. This tragic case inspired me to look at how frequently Ugandan newspapers wrote about child sexual violence. After spending hours in the library searching old newspapers, I came to the realization that this type of violence was extremely common and often reported in the major newspapers. From this point onward, my career path became clear as I developed a fascination for gathering evidence and I often wondered how this knowledge could be used to make a difference in the world.

Upon completion of my undergraduate training, my plans and passion had greatly shifted from teaching history to pursuing a research career, and this guided my search for employment. I was fortunate enough to get my first employment within the field of HIV/AIDS epidemiology and public health with a Uganda-based research institute, Rakia Health Sciences Program (RHSP), a research collaboration between researchers at Makerere University in Uganda, Johns Hopkins University, and Columbia University in the US. While at RHSP, I worked on various research topics on HIV/AIDS and gender-based violence, mainly using qualitative methodology. In the years that followed, I was offered an opportunity to coordinate a team that performed a randomized domestic violence intervention trial (SHARE project). This intervention was successfully implemented and provided me with a wealth of experience and knowledge, and it also made me realize my need for further training.

In 2007 I embarked on a pursuit for knowledge as I undertook my masters training at Linköping University in Sweden. This training offered a research-stimulating environment. I also became more aware of the larger scope of research, and I had opportunities to interact with researchers both among the students and the teaching staff. One crucial interaction was with professor Jan Sundin, one of the professors at the department who took special interest in my master's thesis. In one of our casual fika conversations, we talked about my research goals and I shared with him a bigger picture of where I wanted to end up career wise, and he brought to my attention the idea of pursuing a PhD. After some time, I considered this thought and decided to look for PhD
opportunities with special interest in a position that offered exposure to quantitative methods, because I still felt a gap in that area.

When I started this PhD training, I was fully aware of the challenges that lay ahead. Taking on celiac disease as a focus of my thesis and using a quantitative design wasn’t going to be a walk in the park, especially given my purely non-medical background and bare minimum knowledge of statistics. However, meeting my future team of supervisors during the interview offered a ray of hope. My supervisors were two experienced pediatricians, a statistician and a nutritionist, who all generously pledged their support if I were to be recruited. I had no doubt in my mind when accepting this offer that I would make it through.

During this PhD journey I have taken several quantitative courses and attended a graduate school for register-based research (SINGS). I feel satisfied with the amount of knowledge I have obtained. I am forever grateful for the insights I have gained into the world of celiac disease. Moving to the next step, I still hunger for more knowledge, and I am fully convinced that the researcher in me is already complete although there is a lot of room for growth. I am eager to contribute to the world of science, hopefully joining my efforts with those already in the field, and together we can contribute to making the world a better place.
Acknowledgements

This doctoral project was made possible by a generous grant from the Swedish Research Council, and data were provided by the Swedish Initiative for Microdata Research in the Social and Medical Sciences (Umeå SIMSAM Lab).

To all of you, my colleagues, friends, and family. I say thank you for your moral support during this journey. I would like to mention a few by name.

To my main supervisor Olof Sandström, for generously sharing your excellent knowledge. Your gentleness and patience ensured my success. Above all, thank you for being an excellent listener.

To my co-supervisor and project PI, Anneli Ivarsson, for always showing that you care and for being a role model on how research should be conducted. Your attention to detail is a life lesson that I would like to take with me.

To my co-supervisor Marie Lindkvist, for offering me statistical support and for always boosting my confidence.

To my co-supervisor Cecilia Olsson, your casual yet excellent approach to work made the hardest times of this process a lot easier.

To my co-authors Anna Myleus, Ulf Högberg, Urban Lindgren, and Magnus Strömgren, for generously sharing your expertise.

To each and every one of you at the unit of Epidemiology and Global Health, for creating a great working environment.

To all members of the National Celiac Disease Working Group and those who work with Umeå SIMSAM Lab, for your valuable contribution.

To Fredrik Norström, sharing office space ensured great conversations and knowledge exchange. Thank you for providing great feedback on my thesis.

To Susanne Walther, thank you for your support over the years and for helping with the final touches of my thesis, it was extremely valuable.

To Birgitta Åström and Karin Johansson, for your endless support.

To Evelina Landstedt my officemate for your companionship.
To my neighbors Julia Schröders and Masoud Vaezghasemi, thank you for your generosity.

To Raman Preet, thank you for your unconditional sisterly love and affection. It’s easy to run things by you, you give an honest opinion.

To Elizabeth Höög, thank you for being generous with your time, for your listening ear and never failing ability to calm my nerves.

To Barbara Schumann, for our friendship and for your moral support.

To Ailiana Santosa, thank you for cheering me up.

To my little East African community; Gladys Mahiti, Moses Tetui, Sewe Maquins and Kanyivu Muindi your presence brings a home feeling.

To John Kingsman and Sabina Bergsten, for your friendship.

To professor Jan Sundin, thank you for inspiring me into pursuing a PhD.

To all my colleagues at RHSP. Jennifer Wagman, for your friendship and mentorship. Fred Nalugoda, for your inspiration and valuable friendship. To the SHARE team and ETT, it was a pleasure working with you all.

To my friends Sara Sörensson, Pheobe Berglund and Jackie Bogere, thank you for the laughers and for the ever open doors.

Special thanks go to my family in Uganda and Tanzania, for your endless support and prayers. To my sister Flora, for your unconditional love and sacrifice. To my husband, Leon, for your support over the years. To my children Desire and Joel, you are my world, without you none of this matters.

To my parents who are no longer with us, thank you for teaching me about hard work, commitment, and diligence. My father, at 7 years of age you told me that a PhD was the highest level of education and that I could get it if I wanted. I am sure you would be so proud right now.

Finally, praise to you my God, the source of my strength. Thank you for this victory and for the future that awaits. Psalms 28:7.
References


82. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, van den Brandt PA, Stobberingh EE: Factors influencing the composition


55


pregnant women is different from that of non-pregnant women. 
Appendix

Appendix 1.

The Swedish Working Group for Celiac Disease in Children*

<table>
<thead>
<tr>
<th>Members</th>
<th>The pediatric department representing each regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anneli Ivarsson</td>
<td>Umeå</td>
</tr>
<tr>
<td>Olof Sandström</td>
<td>Umeå</td>
</tr>
<tr>
<td>Niklas P. Nyström</td>
<td>Uppsala</td>
</tr>
<tr>
<td>Ing-Marie Daniels</td>
<td>Uppsala</td>
</tr>
<tr>
<td>Ulrika Fagerberg</td>
<td>Uppsala</td>
</tr>
<tr>
<td>Lars Browaldh</td>
<td>Stockholm</td>
</tr>
<tr>
<td>Lena Grahnquist</td>
<td>Stockholm</td>
</tr>
<tr>
<td>Eva Lindberg</td>
<td>Örebro</td>
</tr>
<tr>
<td>Lars Ekstav</td>
<td>Örebro</td>
</tr>
<tr>
<td>Lotta Högberg</td>
<td>Linköping</td>
</tr>
<tr>
<td>Lars Stenhammar</td>
<td>Linköping</td>
</tr>
<tr>
<td>Audur Gudjonsdottir</td>
<td>Göteborg</td>
</tr>
<tr>
<td>Marianne Malmquist</td>
<td>Göteborg</td>
</tr>
<tr>
<td>Daniel Agardh</td>
<td>Skåne</td>
</tr>
<tr>
<td>Charlotte Neringer</td>
<td>Skåne</td>
</tr>
</tbody>
</table>

*Appointed by the Swedish Pediatric Association through its Section for Gastroenterology, Hepatology, and Nutrition.
Appendix 2.

The National Swedish Childhood Celiac Disease Register: Participating pediatric departments, the pediatrician responsible for reporting the cases to the register, the study are defined by the number of municipalities.

<table>
<thead>
<tr>
<th>City</th>
<th>Reporters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borås</td>
<td>Angelica Steen</td>
</tr>
<tr>
<td>Eskilstuna</td>
<td>Bernadetta Majerczyk</td>
</tr>
<tr>
<td>Falkenberg</td>
<td>Birgitta Davidsson Bårdén</td>
</tr>
<tr>
<td>Falun</td>
<td>Fredrik Quarles van Ufford</td>
</tr>
<tr>
<td>Gällivare</td>
<td>Åke Stenberg</td>
</tr>
<tr>
<td>Gävle</td>
<td>Charlotta Egnell</td>
</tr>
<tr>
<td>Göteborg</td>
<td>Anna Melbin</td>
</tr>
<tr>
<td>Halmstad</td>
<td>Birgitta Bårdén</td>
</tr>
<tr>
<td>Helsingborg</td>
<td>Helen Rosenqvist</td>
</tr>
<tr>
<td>Helsingborg</td>
<td>Monica Lagergren</td>
</tr>
<tr>
<td>Hudinge</td>
<td>Mozaffar Hessami</td>
</tr>
<tr>
<td>Hudiksvall</td>
<td>Ulla Persson</td>
</tr>
<tr>
<td>Härnösand</td>
<td>Tina Melander</td>
</tr>
<tr>
<td>Jönköping</td>
<td>Ulf Jansson</td>
</tr>
<tr>
<td>Jönköping</td>
<td>Inga-Lena Hultman</td>
</tr>
<tr>
<td>Kalmar</td>
<td>Daria Menzel</td>
</tr>
<tr>
<td>Karlskrona</td>
<td>Martin Lindqvist</td>
</tr>
<tr>
<td>Karlstad</td>
<td>Staffan Skogar</td>
</tr>
<tr>
<td>Katrineholm</td>
<td>Bernadetta Majerczyk</td>
</tr>
<tr>
<td>Kristianstad</td>
<td>Roland Schmidt</td>
</tr>
<tr>
<td>Kungsbacka</td>
<td>Anette Ohlsson</td>
</tr>
<tr>
<td>Lidköping</td>
<td>Josefine Hätting</td>
</tr>
<tr>
<td>Linköping</td>
<td>Elisabeth Hollén</td>
</tr>
<tr>
<td>Lund</td>
<td>Charlotte Webb</td>
</tr>
<tr>
<td>Malmö</td>
<td>Daniel Agardh</td>
</tr>
<tr>
<td>Motala</td>
<td>Louise Forslund</td>
</tr>
<tr>
<td>Norrköping</td>
<td>Lotta Högberg</td>
</tr>
<tr>
<td>Norrtälje</td>
<td>Ingmar Zachrisson</td>
</tr>
<tr>
<td>Nyköping</td>
<td>Anna Minkova-Falk</td>
</tr>
<tr>
<td>Skellefteå</td>
<td>Anna Hedlund</td>
</tr>
<tr>
<td>Byrum</td>
<td>Skövde</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Sollefteå</td>
<td>Britta Björsell</td>
</tr>
<tr>
<td>Stockholm</td>
<td>Lars Browaldh (Sachsska)</td>
</tr>
<tr>
<td>Stockholm</td>
<td>- (Astrid Lindgren)</td>
</tr>
<tr>
<td>Sunderbyn</td>
<td>Rune Alenius</td>
</tr>
<tr>
<td>Sundsvall</td>
<td>Jens Bäckström</td>
</tr>
<tr>
<td>Sundsvall</td>
<td>Malin Kristoffersson</td>
</tr>
<tr>
<td>Trollhättan</td>
<td>Nils Wramner</td>
</tr>
<tr>
<td>Uddevalla</td>
<td>Nils Wramner</td>
</tr>
<tr>
<td>Umeå</td>
<td>Olof Sandström</td>
</tr>
<tr>
<td>Uppsala</td>
<td>Niklas Nyström</td>
</tr>
<tr>
<td>Varberg</td>
<td>Anna-Maria Trela</td>
</tr>
<tr>
<td>Visby</td>
<td>Magnus Fredriksson</td>
</tr>
<tr>
<td>Västervik</td>
<td>Jan-Åke Hammersjö</td>
</tr>
<tr>
<td>Västerås</td>
<td>Ulrika Fagerberg</td>
</tr>
<tr>
<td>Växjö</td>
<td>Johan Karlsson</td>
</tr>
<tr>
<td>Ängelholm</td>
<td>Christina Andersson</td>
</tr>
<tr>
<td>Örebro</td>
<td>Eva Lindberg</td>
</tr>
<tr>
<td>Örnsköldsvik</td>
<td>Owe Ljungdahl</td>
</tr>
<tr>
<td>Östersund</td>
<td>Maciej Potyrala</td>
</tr>
</tbody>
</table>
Appendix 3.

The Report form filled in by the reporting pediatrician